



2975.0.55.007 - Census Working Paper 96/4 - Fact Sheet 07 - Income imputation, 1996

Latest ISSUE Released at 11:30 AM (CANBERRA TIME) 01/06/1997

Summary

Main Features

1996 CENSUS OF POPULATION AND HOUSING FACT SHEET 7 INCOME IMPUTATION

IMPUTING A DOLLAR VALUE FOR EACH INCOME RANGE

The 1996 Census was required to produce data on income for persons, families and households. Data was also required for counts of families and households within the various income ranges. The data was collected in ranges, not actual dollars, as this has proven to be the most reliable way to collect income data. The 1996 Census collected the gross weekly income for each person (INCP) using the following income ranges:

range identifier	weekly income
1	Negative
2	income
3	Nil
4	income
5	\$1- \$39
6	\$40 -
7	\$79
8	\$80 -
9	\$119
10	\$120 -
11	\$159
12	\$160 -
13	\$199
14	\$200 -
15	\$299
16	\$300 -
	\$399
	\$400 -
	\$499
	\$500 -

\$599
 \$600 -
 \$699
 \$700 -
 \$799
 \$800 -
 \$999
 \$1,000 -
 \$1,499
 \$1,500
 or more

The collection ranges used on the Census form were chosen after analysing data from the Survey of Income and Housing (SIHC), in which income was collected in actual dollars rather than ranges.

Household and family income

Household and family incomes (HIND and FINF) were derived by summing the personal incomes. However, it is not possible to sum income **ranges**. To overcome this, weighted income data from the SIHC was used to impute a value.

The imputation process

This process involved analysis of the SIHC data to determine the imputation values to be used. Each SIHC record had the appropriate Census range identifier (as defined above) allocated to it. Then, for each range, a mean, median and mid-point were calculated and allocated to each record in that range. Further analysis was needed to determine which of these three measures would be used to impute income for Census records.

Using three 'income groups' - unit income (used in SIHC and not applicable to the Census), family income, and household income, the next step of this exercise analysed the number of groups in each income range when each of the three measures was used, and this was compared with the 'true' number of income groups calculated when using reported income values. For example, when the reported income of individuals was summed to create family income, 10.6% of families had a weekly income of \$120-\$159 (Range 6). When mean values were summed to create family income, slightly more families were in income Range 6 (10.7%). This increased to 10.8% when using median values, and 11.0% when using midpoint values.

This analysis of income units, families, and households was done for Australia, state, and met/ex met regions, and the conclusion was reached that mean values were the most appropriate to use.

For the 1996 Population Census the estimated dollar incomes, calculated using SIHC mean values, were:

range identifier

weekly income

estimated dollar
income

1	Negative income	17
2	Nil income	61
3	\$1 - \$39	99
4	\$40 - \$79	140
5	\$80 - \$119	174
6	\$120 - \$159	244
7	\$160 - \$199	349
8	\$200 - \$299	442
9	\$300 - \$399	540
10	\$400 - \$499	640
11	\$500 - \$599	742
12	\$600 - \$699	873
13	\$700 - \$799	1,163
14	\$800 - \$999	2,650
15	\$1,000 - \$1,499	
16	\$1,500 or more	

These are the values used when summing records to create household and family incomes, and in calculating median values. NOTE: Personal Income is only published in ranges, so these estimated values will not apply.

Problems

There were 3 major limitations inherent in this process:

- (a) calculating median values for open ended ranges; and the related problems of
- (b) the change of income range between personal and household or family incomes; and
- (c) the introduction of sampling error.

1. Median Values for open-ended ranges.

To determine a median value it is necessary to identify the range in which the median lies and then to estimate where within the range the median would be. A problem arises when the median lies within a range which does not have two specified finite end points. After some analysis it was recommended to retain the default value (ie. \$2,000 in the case of HIND and FINF) and indicate, as a table note, that when this value appears in the table the true median income is some value in the range \$2000 or more. (See Footnote (b) below)

2. Change in Income Ranges

The income ranges collected from persons are not the same as the income ranges used for the household and family units. Error is inherent in the calculation of the household and family income ranges due to the imputation of the person level income values. However, since the imputation process did not take into account the change in the income ranges, additional error is introduced due to these changes.

For this census, it was agreed to continue with the different income ranges, given the

apparent difficulty of incorporating any complex changes into Supercross. Moreover, this would be consistent with the process used in the last census. The following paragraph was supplied to Census as an additional note for the family and household income tables to warn of the potential bias:

(a) Due to operational limitations, the family (and household) income imputation methodology may result in an undercount of the number of families (and households) in the \$1500 - \$1999 range and a balancing overcount in the \$2000 range. No other income ranges are affected. This may also affect the median income estimate if the median falls in either of these ranges.

3. Introduction of Sampling Error

The mean income values from the SIHC used as the impute values are subject to sampling error, since SIHC is a sample survey. It would be appropriate to indicate that the household and family income ranges, and therefore any derivations based on these values such as median income values for these units, are subject to both sampling and non-sampling error.

Census agreed to include a suitable statement:

(b) The calculation of median income is based on imputations made from the Survey of Income and Housing Costs, and as such is subject to sampling error. This is particularly evident if the median falls in the highest income range, where the quoted median is a proxy only and should be regarded with caution.

Imputing a dollar value for each income range - Footnotes- to use when creating a HIND or FINF Table

When creating tables containing either Family Income (FINF) or Household Income (HIND) the following issues should be footnotes using the following suggested wording:

(a) Due to operational limitations, the family (household) income imputation methodology may result in an undercount of the number of families/households in the \$1500 - \$1999 range and a balancing overcount in the \$2000 range. No other income ranges are affected. This may also affect the median income estimate if the median falls in either of these ranges.

(b) The calculation of median income is based on imputations made from the Survey of Income and Housing Costs, and as such is subject to sampling error. This is particularly evident if the median falls into the highest income range, where the quoted median is a proxy only and should be regarded with caution.

About this Release

ABOUT THIS RELEASE

Census Fact Sheets are designed to assist in the use and interpretation of Census data. They provide a summary of conceptual and data issues, and changes that have occurred since the last Census. Fact Sheets are developed and produced for a variety of areas, with a focus on a particular question of the Census form. The need for the development of a Fact Sheet is determined by the number of issues raised in a particular area. If there are numerous enquiries, a Fact Sheet will be developed to provide an explanation of the issue.

© Commonwealth of Australia

All data and other material produced by the Australian Bureau of Statistics (ABS) constitutes Commonwealth copyright administered by the ABS. The ABS reserves the right to set out the terms and conditions for the use of such material. Unless otherwise noted, all material on this website – except the ABS logo, the Commonwealth Coat of Arms, and any material protected by a trade mark – is licensed under a Creative Commons Attribution 2.5 Australia licence